BISS Biodiversity Information Science and Standards

Biodiversity Information Science and Standards 10: e183432
doi: 10.3897/biss.10.183432

OPEN ACCESS

Conference Abstract

Fighting Poaching through High-Precision Real-Time Gunshot Detection Using Deep Learning and SAIL

Naveen Dhar ‡

‡ High Tech High Mesa, San Diego, United States of America

Corresponding author: Naveen Dhar (naveendhar8030@gmail.com)

Received: 24 Dec 2025 | Published: 16 Mar 2026

Citation: Dhar N (2026) Fighting Poaching through High-Precision Real-Time Gunshot Detection Using Deep Learning and SAIL. Biodiversity Information Science and Standards 10: e183432.
https://doi.org/10.3897/biss.10.183432

Abstract

Introduction

Passive acoustic monitoring (PAM), thanks to the technological advancement of autonomous recording units (ARUs) (Prince et al. 2019), has enabled long-term, wide-scale studies on natural and/or anthropogenic soundscapes, making it an invaluable tool for conservation (Sugai et al. 2018). Crucially, PAM can aid in detecting threats to biodiversity, such as poaching, by listening for target acoustic signatures. Array-based PAM offers benefits over other poaching monitoring methods, such as unmanned aerial vehicle or camera trap monitoring: lower costs, wider detection fields, and longer-term deployments (Lynam et al. 2025). Real-time detection can enable immediate intervention, making it a goal of the Wildlife Crime Technology Project since 2015 (World Wildlife Fund 2022). However, most acoustic detection of poaching, including existing literature using deep learning for this task, consistently mentions challenges with false positives and employs large fine-tuned models with computational costs better suited for retrospective studies (Katsis et al. 2022, Vu et al. 2024). Our research developed a lightweight deep neural network suitable for on-board, real-time processing and a sensor integration function aiming to reduce false positive rates (FPRs).

© Dhar N. This is an open access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Methods

A pre-partitioned open-source dataset of 35,980 four-second audio recordings from Belizean rainforests (Katsis et al. 2022) was used for training and validation. A test set from Vietnamese rainforests (Vu et al. 2024), comprising 129 background recordings and 19 gunshots, was used for further testing. Waveforms were converted to log-mel spectrograms, an image-like representation of audio displaying the range of frequencies of a signal as it varies with time. Extensive waveform augmentations were applied to the dataset during training to improve the detector's generalizability to new environments. These augmentations included shifting the signal's position in time, injecting background noise, and more.

The lightweight neural network, or model, was designed with convolutional layers and recurrent units for efficient gunshot feature extraction and detection. The final model contained approximately 935k parameters and was trained for 40 epochs.

Many false positives occur because common rainforest sounds resemble gunshots in their frequency-time features (Katsis et al. 2022, Vu et al. 2024). However, gunshot signals propagate farther. Cross-device communication, or the leveraging of information from multiple acoustic sensors, is proposed and investigated as a novel way to reduce the aforementioned false positives in acoustic poaching detection. The proposed Sensor Analysis and Integration Layer, or SAIL function, described in Fig. 1, was developed to integrate predictions from three or more sensors by penalizing isolated confident predictions while boosting sensor agreement.

This was validated through acoustic propagation simulations on the Vietnamese dataset, where each waveform had three copies with random augmentations mimicking distance-dependent attenuation, emulating a three-sensor array. The model generated confidence scores for each copy, which were then fed into SAIL to create a final probability. Across the 1000 simulation runs performed on the dataset, Wilcoxon signed-rank tests evaluated whether median differences in mean FPR per-file exceeded zero.

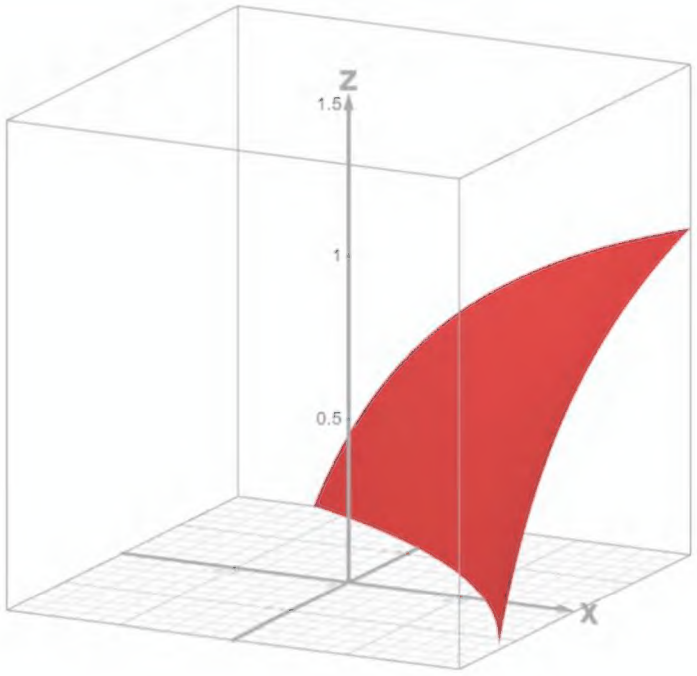
Results

Model performance was evaluated on the validation partition of the Belizean dataset. A precision-recall curve (Fig. 2) showed strong generalization and 0.96 average precision. Table 1 compares the lightweight model against the leading open-source rainforest gunshot detector, or the fine-tuned ResNet18 from Katsis et al. (2022). Our lightweight model is 11MB and compresses to 964KB via TensorflowLite serialization. Table 2 summarizes SAIL simulation statistics, reporting mean FPR differences and Wilcoxon p-values alongside performance metrics. SAIL produced significant FPR reduction with minimal (1%) recall reduction.

Conclusion

While not directly extendable to efficacy in the field, statistical evaluation demonstrates that SAIL consistently reduces FPR compared to simple inference across a wide variety of

simulated sensor arrays. The lightweight gunshot detector generalizes to the validation dataset at 935k parameters, retaining 94% of performance (validation F1 at 95% recall) at a fraction of the leading work's model size. Overall, efficient deep learning and SAIL present a promising approach towards cost-effective real-time poaching detection.



$$P_f(p_1, \ldots, p_n) = \sigma\left(1 - \frac{\sum_{1 \le i < j \le n} (1 - p_i)(1 - p_j)}{\sum_{k=1}^{n} p_k}\right)$$

$$\sigma(x) = \frac{1}{1 + e^{k(x-m)}}$$

Figure 1.

A 3D graph of the SAIL function with n=3, k=10, and m=0.5. The x and y values (sensor predictions p1 and p2) are constrained from 0 to 1. In order to visualize the function, p3 was fixed with a value of 0.5.

"Sensor Analysis and Integration Layer (SAIL) Graph and Equation" by Naveen Dhar is licensed under CC BY 4.0

Table 1.

Comparing the performance and model file size of the leading deep-learning classifier for gunshot detection on the Belizean dataset with the best-performing model developed in this study. F1 is the harmonic mean between precision and recall. Distinct F1 denotes the best F1 on the spatially distinct Vietnamese dataset.

Classifier	F1 (Best)	F1 @ 95% Recall	F1 (Distinct)	Model Size	Parameters
Katsis et al. (ResNet18)	Unknown	0.89	Unknown	87 MB	≈ 11.7 M
Proposed	0.91	0.85	0.77	11 MB	935 K

Table 2.

Summary of Wilcoxon statistical testing of SAIL simulation alongside performance metrics, comparing inference and SAIL across the 1000 simulations.

Metric	SAIL Mean (Per-run average)	Inference Mean (Per-run average)	Mean Per-File Difference (Inference - SAIL)	Wilcoxon p-value
FPR Difference (Inference - SAIL)	-	-	0.024	0.04
Recall	0.616	0.684	-	-
F1	0.762	0.743	-	-
Precision	1.000	0.813	-	-

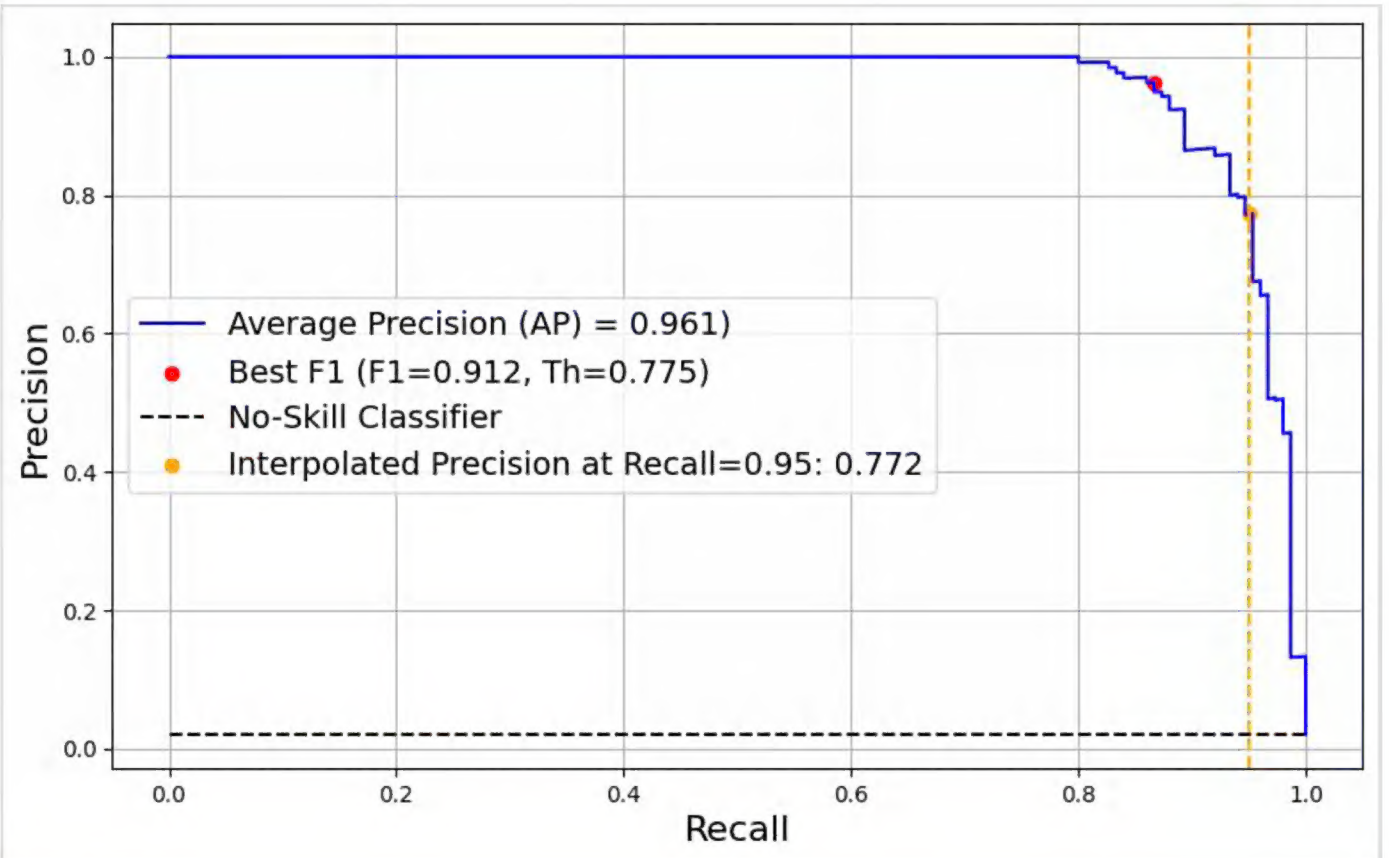


Figure 2.

Precision-recall curve of the lightweight model on the validation partition of the Belizean dataset.

"Lightweight Rainforest Gunshot Detection Network PR-Curve" by Naveen Dhar is licensed under CC BY 4.0

Keywords

passive acoustic monitoring, AI on the edge, IoT, poaching detection, autonomous recording units, rainforest conservation

Presenting author

Naveen Dhar

Presented at

Living Data 2025

Conflicts of interest

The authors have declared that no competing interests exist.

References

- Katsis LD, Hill A, Piña-Covarrubias Eao (2022) Automated detection of gunshots in tropical forests using convolutional neural networks. Ecological Indicators 141 https://doi.org/10.1016/j.ecolind.2022.109128
- Lynam AJ, Cronin D, Wich S, Steward J, Howe Aao (2025) The rising tide of conservation technology: empowering the fight against poaching and unsustainable wildlife harvest. Frontiers in Ecology and Evolution 13 https://doi.org/10.3389/fevo.2025.1527976
- Prince P, Hill A, Piña Covarrubias E, Doncaster P, Snaddon J, Rogers A, et al. (2019) Deploying Acoustic Detection Algorithms on Low-Cost, Open-Source Acoustic Sensors for Environmental Monitoring. Sensors 19 (3). https://doi.org/10.3390/s19030553
- Sugai LSM, Silva TSF, Ribeiro JW, Llusia D, et al. (2018) Terrestrial Passive Acoustic Monitoring: Review and Perspectives. BioScience 69 (1): 15 -25. https://doi.org/10.1093/biosci/biy147
- Vu TT, Phan DV, Le TS, Clink DJ (2024) Investigating hunting in a protected area in Southeast Asia using passive acoustic monitoring with mobile smartphones and deep learning. Ecological Indicators 167 https://doi.org/10.1016/j.ecolind.2024.112501
- World Wildlife Fund (2022) Wildlife Crime Technology Project. https://www.worldwildlife.org/our-work/wildlife/wildlife-crime/wildlife-crime-technology-project/